Synopsis of the Ph.D. thesis entitled

EVALUATING THE FEASIBILITY OF CONSIDERATE VALUATION MODEL FOR REAL ESTATE PROPERTIES IN AURANGABAD CITY

Submitted to

Faculty of Engineering and Technology,

MGM University, Chhatrapati Sambhajinagar (MS) - 431003.

By

NANDKISHOR BALAJI SONAR

(PRN – 202101119004)

For the award of Ph. D. degree

DOCTOR OF PHILOSOPHY IN THE SUBJECT OF CIVIL ENGINEERING

Under the guidance of

Dr. Vijaya Sitaram Pradhan

Professor,
Civil Engineering Department,
Jawaharlal Nehru Engineering College,
MGM University, Chhatrapati Sambhajinagar (MS) - 431003.

January 2025

Index

Sr. no.	Content	Page no.
1	Introduction	1
2	Objectives	2
3	Scope of work	2
4	Work done	2
5	Contributions made to the existing knowledge in the subject	3
6	Data collection and analysis	5
7	Results and discussion	7
8	Conclusions	12
9	Bibliography	13
10	List of publications	14

List of tables

Sr. no.	Content	Page no.
1	Rank wise list of value indicators from Questionnaire survey	5
2	List of value indicators from Pairwise comparison matrix	6
3	Results of rank correlation	7
4	First ten ranks of value indicators	8
5	Final ten crucial value indicators	8
6	Regression parameters for 90-10 split	9
7	Accuracy of LR, RFR and GBR models across all four splits	11

List of figures

Sr. no.	Content	Page no.
1	Comparison of actual and predicted values in 80-20 split of 2020 for LR	10
2	Scatter plot in 80-20 split of 2020 for LR	10
3	LR model’s accuracy over the years and across various splits	11

Introduction

Property valuation is a complex and time-consuming process which is carried out by qualified real estate appraisers. Number of properties and number of purchase-sale transactions grows year by year. Mass real estate appraisal appears as another big problem, connected with deficiency of human and time resources. Therefore, numerous studies are carried out on computer systems which can support the real estate appraisers. (Malinowski et. al., 2018)

The emerging techniques applied in the field of real estate valuation includes hedonic pricing method, regression analysis or a combination of computer application and a specific framework with independent parameters to determine the dependent property value. Such tools use available data, analyse its pattern and provide output with the help of computer algorithm. Multiple linear regression is commonly used to predict property values based on various influencing factors. Understanding its principles, assumptions and evaluation metrics is crucial for interpreting results and making informed decisions based on the analysis.

Decision-making is a critical part of a typical real estate property valuation aimed at quantifying market value of a property according to its qualitative characteristics. A new approach that regards decision making as a data-driven process whose tasks can be decomposed and made executable as granular applications individually can be implemented. (Michele Argiolas et al., 2010) Application of hedonic pricing method and regression analysis may serve the purpose. Machine learning may be a supplementary input in this which may make valuation process simpler than it already is.

The studied literature collectively emphasizes the need of hedonic pricing models and machine learning techniques to accurately understand housing prices. Integration of spatial characteristics and advanced predictive methods signifies a comprehensive approach to property valuation. As the complexity of real estate markets continues to evolve, these studies advocate for the ongoing refinement and integration of innovative methodologies that enhance the accuracy of real estate property valuation. Therefore, this research work concentrates on implementation of machine learning techniques to predict the property values. For this purpose, relevant regression models like Linear Regression, Random Forest Regressor and Gradient Boosting Regressor were used to compare their suitability. Thereby, the objective was to determine the best suited model for a varied dataset along with the most compatible training-testing split based on regression metrics and prediction accuracy.

Objectives

1. To determine people's perception about the crucial indicators of property value by conducting questionnaire survey, personal interviews of valuers, pairwise comparison matrix and their correlation analysis.
2. To develop a valuation framework of multiple linear regression using machine learning techniques and evaluate its performance by studying its regression parameters, prediction accuracy, compatibility of various training-testing splits and cross-validation.
3. Compare the property values calculated by the above-mentioned framework with the documented values as observed in the sales transactions of government records.

Scope of work

This study focuses on identifying the prominent value indicators of a real estate property, collecting sales transactions of properties in the study area, developing a regression model and implementing it to determine the fair market values of all these properties. The functionality of the model will be verified through its regression parameters and prediction accuracy. Finally, the calculated values will be compared with the documented values of sales transactions.

Questionnaire survey, valuer's interviews and pairwise comparison matrix will be implemented to identify crucial value indicators, which will be framed through similar research papers and by consulting valuers. Correlation analysis of above-mentioned approaches will be performed using Kendall's and Spearman's rank correlation coefficients, to shortlist crucial value indicators. A regression-based valuation framework with machine learning techniques namely – Linear Regression, Random Forest Regressor and Gradient Boosting Regressor along with different sets of training-testing splits (90-10, 80-20, 70-30, 60-40) will then be created to determine property values for the study area – Aurangabad (formally known as Chhatrapati Sambhajinagar) over the years 2020-2022, which will be validated from government records.

Work done

- Selecting 34 value indicators from literature, valuers, home buyers and categorizing them. Conducting Questionnaire survey (QS) of general public, Personal interviews (PI) with practicing valuers, Pairwise comparison matrix (PCM). Identifying the most crucial value

indicators based on the results of above three approaches as well as their Correlation analysis using Kendall's and Spearman's rank correlation methods.

- Collecting more than fifteen thousand sales transactions (Index 2 sheets with details of property transaction date, market value in rupees as per government records, area in square meters, location, type of property i.e. residential (flat, row-house, bungalow) / commercial (shop, office) / open land) from the sub-registrar's office, Aurangabad, Maharashtra, India for the years 2020 to 2022. To ensure data quality, discarding repeated entries, gift deeds, mortgage deeds and correction letters etc., which do not accurately reflect fair market values. Additionally, transactions with abnormally low or high values were removed. The dataset was structured in a clean, tabular format with columns representing the input features and target variable.
- Selecting regression models namely – Linear Regression (LR), Random Forest Regressor (RFR) and Gradient Boosting Regressor (GBR) based on their suitability for the dataset. Selecting value indicators compatible with the regression model for further analysis. Implementing different training-testing splits (e.g. 90% training and 10% testing, 80% training and 20% testing, 70% training and 30% testing, 60% training and 40% testing) to ensure model's robustness. Regression analysis of the dataset in a Python environment using Google Colab, incorporating libraries such as scikit-learn for implementation in the specific algorithm of each model.
- Calculating regression metrics like coefficient of correlation, coefficient of determination (R^2), average error, standard error, mean absolute error (MAE), root mean squared error (RMSE) for 2020, 2021, 2022 and in a consolidated manner for all three years 2020-2022. Five-fold cross-validation of the regression results in all three models. Evaluating each model's performance using regression summary and accuracy percentages. Comparing predicted values with the actual ones using the model's evaluation metrics.

Contributions made to the existing knowledge in the subject

Novel findings: The thesis will provide a new technique of selecting crucial value indicators and empirical analysis of existing sales transactions to predict property values of proposed real estate properties, by three distinct approaches along with their correlation, four categorised training-testing splits and three different regression models using specific algorithms.

Addressing knowledge gaps: This research will fill the gap in existing literature and provide new insights in the under-studied fields of real estate valuation.

Practical applications: The research will guide in the application of regression models for determining property values as a solution of mass property appraisal problems.

Critical analysis: The thesis will critically evaluate existing work, theories, methods or practices in real estate valuation, leading to deeper knowledge gain of the related processes.

Knowledge integration: The research work integrates comprehensive information from the disciplines of civil engineering, valuation, statistical as well as regression analysis, modelling.

Policy implications: The findings will influence decision makers about framing policies of standardised valuation system.

IPR development: This research led to the grant of a patent, two conference presentations with published proceedings and five paper publications in UGC indexed journal / peer reviewed unpaid journal / SCOPUS indexed journal.

Originality: It can be noticed from application of Pairwise comparison approach in determining value indicators and its further comparison with other established approaches. This study has enough potential to select appropriate value indicators, its preferential order, its quantification and its application for determining property value. Also, regression equation was generated in the LR model with understudied independent variables:

X1 - Market value of the property (as per government records, in Rupees)

X2 - Locality of the property

X3 - Detailed address of the property

X4 - Type of property (residential/commercial/open land)

X5 - Floor number

X6 - Saleable area (in square meters)

X7 - Transaction date.

Implementing above variables, following regression equation (1) for property value ('Y' in rupees) was framed,

$$Y = \beta 0 + \beta 1 * X1 + \beta 2 * X2 + \beta 3 * X3 + \beta 4 * X4 + \beta 5 * X5 + \beta 6 * X6 + \beta 7 * X7 \qquad (1)$$

here, $\beta 0$ = Y-intercept of regression line, $\beta 1$ - $\beta 7$ = coefficients of respective variables *X1* - *X7*.

Research Limitations: Limitations of this study includes the selected independent variables on the basis of which dependent variable's value is determined.

Data collection and analysis

QS was conducted using Google forms circulated through emails and WhatsApp service to 1000 individuals, out of which 343 responses were received. The responses to all 34 questions related to each value indicator were collected in the form of likert scale rating between 1 to 5 where, 1 denoted 'very low' rating, 2 denoted 'low', 3 denoted 'moderate', 4 denoted 'high' and 5 denoted 'very high' rating. The value indicators were ranked according to their weighted mean in the descending order. Wherever a tie was observed in the ranks, the average of those ranks was allotted to the ties. The most frequent response was also calculated for each question along with the range. It showed that the responses had relatively high variability with the range of 4 for each question. The relatively high variability among the respondents suggests moderate agreement for each of the value indicators. Following table 1 shows questions as per their category and rank with respect to weighted mean. Similar table is also prepared for PI analysis.

Table 1 – Rank wise list of value indicators from QS

Value indicator	Most frequent response	Mean	Rank	Category	SD
Continuous electric supply	5	4.4519	1	Service facility	0.8098
Regular water supply	5	4.3848	2	Service facility	0.8489
Locality around	5	4.3557	3	Location details	0.7947
Good ventilation	5	4.2770	4	Building details	0.8651
Personal parking facility	5	4.2741	5	Service facility	0.8193
Availability of lift	5	4.1720	6	Service facility	0.9520
Near transport	5	4.1633	7	Location details	0.9428
Construction quality	4	4.1341	8.5	Building details	0.8603
Legal control of property	5	4.1341	8.5	Building details	0.9097
Returns from property	5	4.1254	10	Building details	0.8996
External view	5	4.1224	11	Building details	0.8886
Security services	5	4.0379	12	Service facility	0.9846
Age of building	4	4.0292	13	Building details	0.8290
Size of property	5	3.9942	14	Building details	0.9598
Public transport	4	3.9679	15	Location details	0.9300
Free access to property	4	3.9417	16	Building details	0.9550
Particular floor	4	3.9359	17	Building details	0.8980
Near market	4	3.9184	18	Location details	1.0241
Total rooms	4	3.8980	19	Building details	0.9058
Road conditions	4	3.8950	20	Location details	0.9871
Future expansion	4	3.8426	21.5	Building details	1.0036
Total bedrooms	4	3.8426	21.5	Building details	0.9591
Near workplace	4	3.7988	23	Location details	0.9370
Insurance of property	4	3.7959	24	Building details	1.0438

Solar panels	4	3.7843	25	Building details	1.0691
Depreciation	4	3.7580	26	Building details	1.0258
Rain water harvesting	4	3.7522	28	Building details	1.0798
Total bathrooms	4	3.7522	28	Building details	0.9226
Maintenance charges	4	3.7522	28	Building details	0.9382
Vastu shastra	4	3.7172	30	Building details	1.1139
Total floors	4	3.6968	31	Building details	0.9605
Taxes	4	3.6735	32	Building details	1.0322
Groundwater	4	3.6531	33	Location details	1.0687
Pollution in area	3	3.6385	34	Location details	1.1134

To accomplish the task of ascertaining value indicators of a property, personal visits were made to more than 50 valuation professionals. However, only 34 out of them agreed to respond. Its analysis was done in the similar manner as QS responses.

In the pairwise comparison matrix approach, all 34 selected value indicators were compared face to face in a 34 x 34 block matrix for calculating sum scores of each value indicator. Table 2 below shows the rank wise list on the basis of this sum score.

Table 2 – List of value indicators from PCM

Value indicator	Sum score	Rank	Category
Continuous electric supply	33	1	Service facility
Regular water supply	32	2	Service facility
Locality around	31	3	Location details
Construction quality	28.5	4	Building details
Near transport	27	5.5	Location details
Size of property	27	5.5	Building details
Personal parking facility	25	7.5	Service facility
Security services	25	7.5	Service facility
Near market	24	9.5	Location details
Public transport	24	9.5	Location details
Legal control of property	23	11.5	Building details
Availability of lift	23	11.5	Service facility
Returns from property	22	13	Building details
Free access to property	20	14	Building details
Good ventilation	19	16	Building details
External view	19	16	Building details
Particular floor	19	16	Building details
Age of building	18.5	18	Building details
Total rooms	15	19	Building details
Road conditions	14	20	Location details
Future expansion	12.5	21.5	Building details
Total bedrooms	12.5	21.5	Building details
Near workplace	11	23	Location details

Insurance of property	10	24	Building details
Solar panels	9	25.5	Building details
Depreciation	9	25.5	Building details
Rain water harvesting	6	28	Building details
Total bathrooms	6	28	Building details
Maintenance charges	6	28	Building details
Vastu shastra	4	30	Building details
Total floors	3	31	Building details
Taxes	2	32	Building details
Groundwater	1	33	Location details
Pollution in area	0	34	Location details

The correlation analysis among the ranks of QS, PI and PCM was done with each other in three sets of each pair, first using Kendall's rank correlation coefficient and then Spearman's rank correlation coefficient. The results of comparison of all the three approaches using the two correlation methods were tabulated and based on that, conclusions regarding the most significant correlation method and approach were framed.

Results and discussion

Since value indicators play a vital role in the valuation process, they need to be cautiously identified. The adopted three approaches guided in the selection of such crucial indicators. It was noted that, correlation analysis by Kendall's method underscored the comparison as lowest whereas, Spearman's method as highest for all the three approaches. It is evident from the correlation coefficient of 0.96 that QS and PCM were closely related, which was calculated by the Spearman's method. Following table 3 highlights the values of rank correlation coefficients.

Table 3 – Results of rank correlation

Approach	**Methods**	
	Kendall's	**Spearman's**
Valuer's responses and Questionnaire survey	0.46	0.70
Valuer's responses and Pairwise comparison matrix	0.57	0.79
Pairwise comparison matrix and Questionnaire survey	0.84	0.96

Since Spearman's rank correlation coefficient of all three approaches is highest, between 0.70 to 0.96, it signifies that the three approaches are strongly correlated. The first ten ranks of value indicators obtained from QS, PI and PCM were listed as shown in table 4 below to facilitate in the shortlisting of the most influential ones. First preference was given to the common indicators in all three approaches and then to indicators that were common in any of the two

approaches. For example – Locality around is common in all three approaches so it was given first preference whereas, Size of property is common in two of the three approaches so it was given second preference and so on for other indicators from the list.

Table 4 – First ten ranks of value indicators

Rank	Valuer's responses	Questionnaire survey	Pairwise comp. matrix
1	Size of property	Cont. electric supply	Cont. electric supply
2	Locality around	Regular water supply	Regular water supply
3	Near market	Locality around	Locality around
4	Construction quality	Good ventilation	Construction quality
5	Public transport	Personal parking facility	Near transport
6	Regular water supply	Availability of lift	Size of property
7	Continuous electric supply	Near transport	Personal parking facility
8	Particular floor	Construction quality	Security services
9	Availability of lift	Legal control of property	Near Market
10	Maintenance charges	Returns from property	Public transport

Table 5 below highlights the final ten crucial value indicators from the three approaches along with their category, which are shortlisted from table 4.

Table 5 – Final ten crucial value indicators

Approach	Value indicators	Category
PI, QS, PCM	Continuous electric supply	Service facility
	Regular water supply	Service facility
	Locality around	Location details
	Construction quality	Building details
QS and PCM	Personal parking facility	Service facility
	Near transport	Location details
PI and PCM	Size of property	Building details
	Near Market	Location details
	Public transport	Location details
PI and QS	Availability of lift	Service facility

A dataset of 4541 sales transactions was collected from the sub-registrar's office for the years 2020 to 2022, comprising 1991 transactions in 2020, 1784 transactions in 2021 and 766 transactions in 2022. Linear Regression, Random Forest Regressor and Gradient Boosting Regressor models were trained and tested with this dataset. The data analysis evaluated performance of these models, across four training-test splits (90-10, 80-20, 70-30 and 60-40) over the years 2020, 2021, 2022 and a combined analysis for 2020-2022. These performance metrics for all the mentioned splits are presented in the sample table 6 below. RMSE and R-squared measure the goodness of fit of a linear regression model to a dataset. RMSE indicates

the model's ability to predict the value of the dependent variable, while R-squared reflects how well the independent variables account for the variation in the dependent variable. When comparing the accuracy of different linear regression models, RMSE is generally a more suitable choice than R-squared. Similar tables were also made for other set of splits i.e. 80-20, 70-30 and 60-40.

Table 6 – Regression parameters for 90-10 split

90-10	Model	Multiple R	R-Squared	Standard Error	Mean Absolute Error	Root Mean Squared Error	Average Error
2020	LR	0.9400	0.8836	273380.73	519698.73	854049.28	3.816304
	RFR	0.9400	0.8836	301780.20	576897.27	1378709.77	7.924701
	GBR	0.9400	0.8836	284872.54	592775.37	1401184.38	5.761224
2021	LR	0.9477	0.8982	717641.35	718049.44	1187498.32	3.002955
	RFR	0.9477	0.8982	483831.81	1141932.65	4063922.81	5.335226
	GBR	0.9477	0.8982	533152.23	979929.97	3020200.49	4.569951
2022	LR	0.8900	0.7921	564528.36	617412.17	889842.72	4.217461
	RFR	0.8900	0.7921	584577.63	700946.38	1383061.68	4.158566
	GBR	0.8900	0.7921	556790.17	649802.58	1079759.92	2.219371
2020 - 2022	LR	0.9490	0.9007	171507.26	853159.78	2438584.23	7.162358
	RFR	0.9490	0.9007	187708.94	768617.16	2408996.05	4.234515
	GBR	0.9490	0.9007	182522.17	759715.43	2459092.20	5.777331

Overall, LR model consistently showed robust performance across different splits and years, though it experienced ups and downs in error metrics. Alternatively, its average RMSE and average error were the lowest which shows that it was the most compatible model for determining the real estate property values. RFR generally underperforms compared to LR. In certain years, it showed significant spikes in Mean Absolute Error, particularly in 2021. GBR's performance lagged behind in most cases but offered competitiveness in certain splits, highlighting its adaptability. Also, the low values of average error signify that the models have performed well. All these models appear sensitive to changes in data distribution or market conditions, as evidenced by the increased error metrics observed in 2021 and 2022.

LR model generated regression equations for all splits and across all study years, comprising of all seven variables and provided them in the following sample form for 90-10 split of 2020:

$$y = 360211.09 + (1.14 * X1) + (554.78 * X2) + (53.26 * X3) + (48648.03 * X4) + (14765.51 * X5) + (2328.22 * X6) + (-0.00 * X7) \quad (2)$$

Notably, *X1* (market value) was the most significant determinant of property value. Also, *X2* (locality) and *X4* (property type) prominently influenced valuation. Whereas, *X3* (detailed addresses), *X5* (floor number) and *X6* (saleable area) carried mixed, context-sensitive outcomes. The coefficients for *X7* (transaction date) revealed minimal impact on *y* (property value), suggesting that its direct influence on valuation, based on the transaction date is minimal. Figures 1 below shows sample graph of difference between actual and predicted values whereas, figure 2 shows the scatter plot for LR in 80-20 split, along with regression line. Similar year wise graphs were also prepared for all other models and combinations of splits.

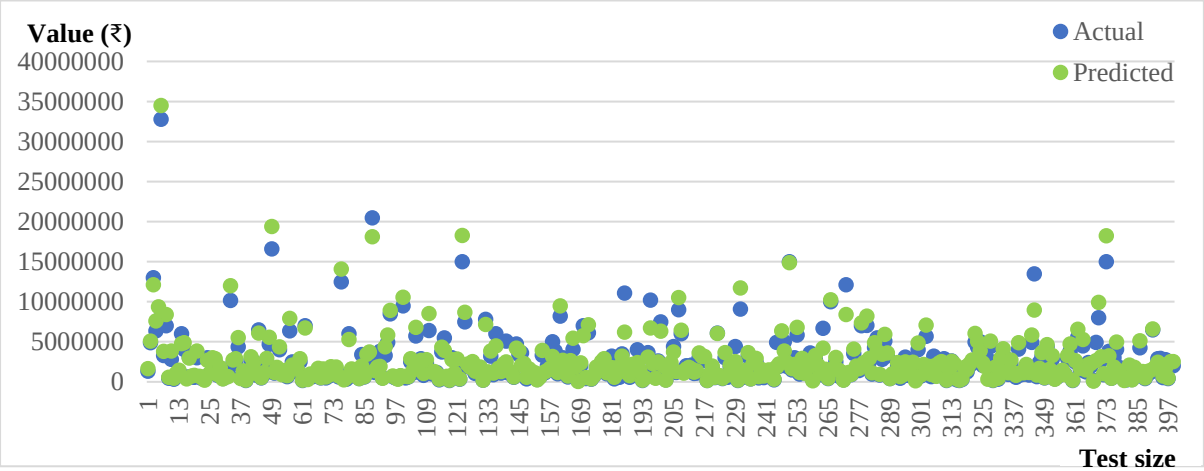

Figure 1 – Comparison of actual and predicted values in 80-20 split of 2020 for LR

It is distinct from the low difference between actual and predicted values that they are very close to each other. Hence reinforcing the robust performance of all the models.

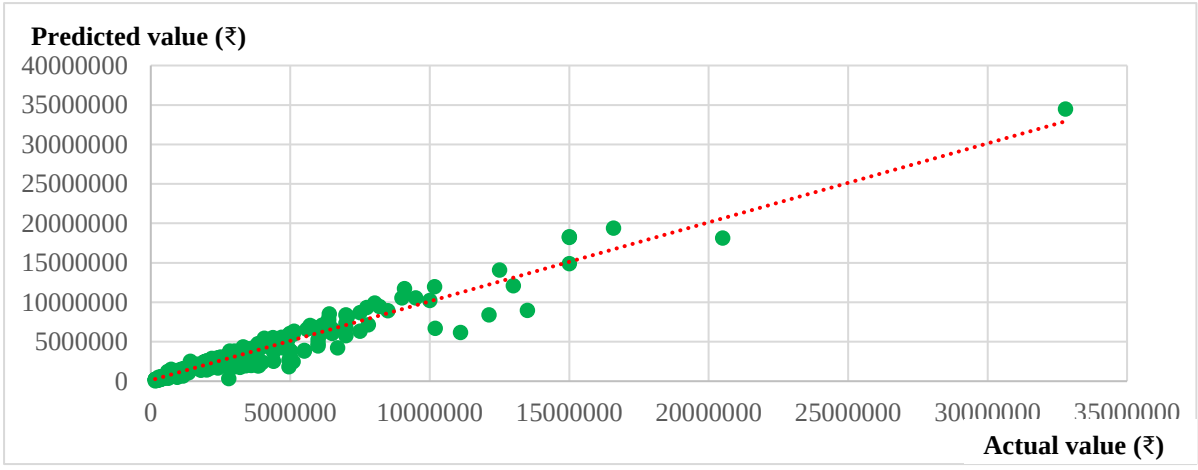

Figure 2 – Scatter plot in 80-20 split of 2020 for LR

These scatter plots clearly show that the intersection of actual and predicted values lie very close to the regression line which highlights the high efficiency of all three models.

LR model showed high accuracy across all years and splits, affirming its effectiveness. RFR exhibited variable performance, indicating potential issues in overfitting. GBR performed relatively well with larger data sizes, indicating good adaptive learning capabilities. Across all models, 90-10 split consistently yielded highest accuracy, demonstrating that increased training data enhances model's performance. Table 7 and sample figure 3 help to visualize this analysis. Similar graphs were also prepared for other models – RFR, GBR.

Table 7 Accuracy of LR, RFR and GBR models across all four splits

Accuracy (%)		90-10	80-20	70-30	60-40
LR	2020	81.25	80.41	79.85	78.09
	2021	83.74	81.9	80.84	80.33
	2022	83.93	80.22	76.06	76.58
	2020-2022	75.34	76.92	78.14	78.3
Accuracy (%)		**90-10**	**80-20**	**70-30**	**60-40**
RFR	2020	79.18	81.22	78.98	78.4
	2021	74.14	77.89	78.47	78.27
	2022	81.75	77.92	76.36	77.88
	2020-2022	77.79	79.78	80.73	81.2
Accuracy (%)		**90-10**	**80-20**	**70-30**	**60-40**
GBR	2020	78.61	80.38	78.53	77.65
	2021	77.81	79.11	80.18	78.45
	2022	83.09	75.53	76.26	77.76
	2020-2022	78.04	79.91	80.53	81.2

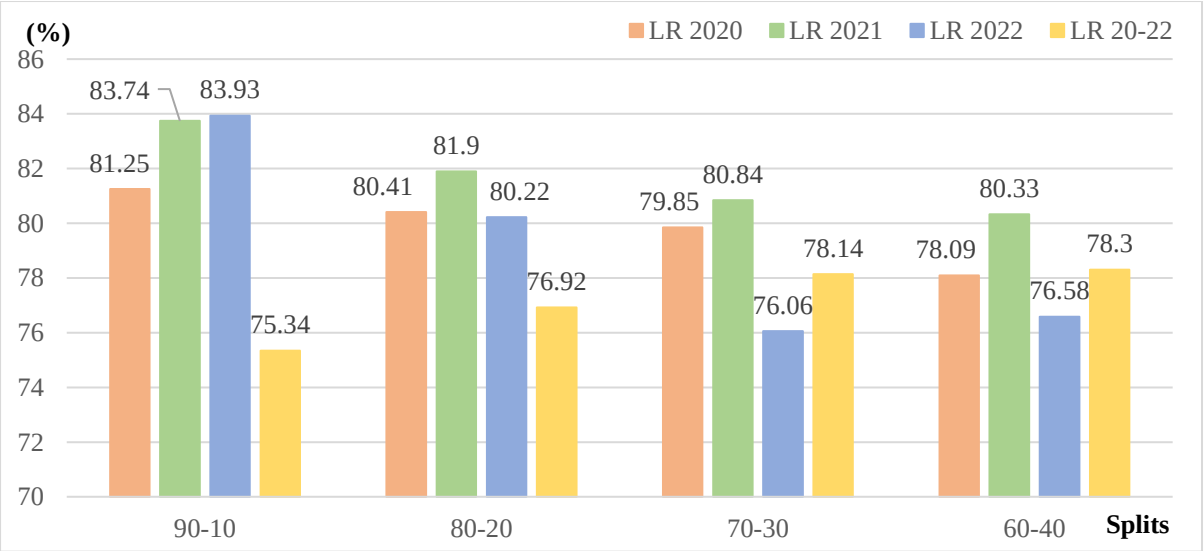

Figure 3 – LR model's accuracy over the years and across various splits

The LR model's accuracy ranged from 75.34% (2020-2022, maximum dataset) to 83.93% (2022, minimum dataset). The RFR model's accuracy ranged from 74.14% (2021)

to 81.75% (2022). This model performed relatively well in 2020 but did not reach higher accuracy levels compared to LR. The GBR model's accuracy ranged from 75.53% (2022) to 83.09% (2022) showing a competitive capacity and similar pattern to RFR.

Conclusions

- The respondent's perception about the crucial value indicators of a property was collected by adopting three approaches – QS, PI, PCM which guided in the selection of such value indicators. It was noted that, correlation analysis gave satisfactory results for all these three approaches based on which, ten crucial value indicators were shortlisted.
- Comprehensive analysis of developed valuation framework and its performance was evaluated using regression models – LR, RFR and GBR, which was conducted across various training-test splits (90-10, 80-20, 70-30, and 60-40) from 2020 to 2022 yielding a plethora of insights into their predictive capabilities. The key performance metrics such as Multiple R, R^2, Average Error, Standard Error, MAE and RMSE allowed for informed comparisons. The results emphasized that, LR model achieved the best performance across this varied dataset, exhibiting lowest RMSE and the highest accuracy of 83.93%, particularly with 90-10 training-testing split. In contrast, RFR and GBR required careful validation and adaptability to provide similar levels of accuracy and reliability. Thus, this analysis highlights robustness of LR against complex market trends, whereas RFR and GBR exhibit variability in adapting to changing data.
- Ultimately through this analysis, it is evident that an optimal data split can significantly influence the model's predictive capabilities. The 90-10 split generally outperformed others across most metrics and accuracy levels, suggesting that a larger training dataset enhances model's performance. In contrast, smaller splits (80-20, 70-30, and 60-40) showed a tendency to reveal higher error metrics and less stable predictions, particularly for RFR and GBR, which highlighted their sensitivity to available training data.
- Additionally, the model's predictive performance was most effective when applied to the comprehensive dataset spanning all three years, as evidenced by its impressive R-squared values and minimal RMSE across all splits. However, the model faced challenges in 2022, which had the least amount of data available. A notably consistent coefficient of determination, approximately 0.87, was recorded in the 70-30 split, demonstrating that the selected seven independent variables significantly and collectively contributed to explain the variations in the dependent variable.

- The comparison of predicted values from the models was done with documented values of government records based on the output of testing split and percentage accuracy of all three models. It is evident from high accuracy values of all models that the predicted values are very close to the documented values, justifying the application of the implemented models.

Bibliography

- Adrian Malinowski, Mateusz Piwowarczyk, Zbigniew Telec, Bogdan Trawinski, Olgierd Kempa, & Tadeusz Lasota. (2018). An approach to property valuation based on market segmentation with crisp and fuzzy clustering. In N. T. Nguyen et al. (Eds.), *ICCCI 2018* (LNAI 11055, pp. 534–548). Springer Nature Switzerland AG.
- Firano Zakaria & Filali A. Fatine. (2021). Towards the hedonic modeling and determinants of real estate price in Morocco. *Journal of Social Sciences and Humanities Open, 4*, Article 100176 (pp. 01-09).
- Gongzhu Hu, Jinping Wang, & Wenying Feng. (2013). Multivariate regression modeling for home value estimates with evaluation using maximum information coefficient. In R. Lee (Ed.), *Software Engineering, Artificial Intelligence, Networking* (SCI 443, pp. 69–81). Springer-Verlag Berlin Heidelberg.
- Juan R. R. & Paloma Taltavull De La Paz. (2021). Machine learning with explainability or spatial hedonics tools? An analysis of the asking prices in the housing market in Alicante, Spain. *Journal of Expert Systems With Applications, 171*, Article 114590 (pp. 01-14).
- Krzysztof Olszewski, Joanna Waszczuk, & Marta Widlak. (2017). Spatial and hedonic analysis of house price dynamics in Warsaw, Poland. *Journal of Urban Planning and Development, 143*(3), Article 04017009 (pp. 01-10).
- Michele Argiolas, Nicoletta Dessi, Giampaolo Marchi, & Barbara Pes. (2010). Real estate decision making processes and web-based applications: An integrated approach. In G. Phillips-Wren et al. (Eds.), *Advances in intelligent decision technologies* (SIST 4, pp. 329–338). Springer-Verlag Berlin Heidelberg.
- Omotayo O. O., Arum C. & Ikumapayi C. M. (2024). Assessment of machine learning methods for concrete compressive strength prediction. *Journal of Soft Computing in Civil Engineering, 8*(4), pp. 01–06.
- Sayali Sandbhor & N. B. Chaphalkar. (2019). Effect of training sample and network characteristics in neural network-based real property value prediction. In A. J. Kulkarni et al. (Eds.), *Proceedings of the 2nd International Conference on Data Engineering and*

Communication Technology (Advances in Intelligent Systems and Computing 828). Springer Nature Singapore Pte. Ltd., pp. 303-313.

- Shanaka Herath. (2013). Discovering the urban structure using a spatial hedonic house price model. *Journal of Smart Construction and Management in the Context of New Technology (ICCREM),* ASCE, pp. 1352-1372.
- Stephen Wakaba Gatheru, David Nyika. (2015). Application Of Geographic Information System In Property Valuation. *International Journal of Scientific & Technology Research 4*(8), ISSN 2277-8616, pp. 61-71.
- Sukran Yalpir, Suleyman Sisman, Ali Utku Akar, & Fatma Bunyan Unel. (2021). Feature selection applications and model validation for mass real estate valuation systems. *Journal of Land Use Policy, 108,* Article 105539, pp. 01-18.
- Tadeusz Lasota, Tomasz Luczak, & Bogdan Trawinski. (2011). Investigation of random subspace and random forest methods applied to property valuation data. In P. Jedrzejowicz et al. (Eds.), *ICCCI 2011* (Part I, LNCS 6922, pp. 142–151). Springer-Verlag Berlin Heidel.
- Yalei Wu, Junjie Yang, Xiaoli Liu, Yuting Lu, & Ruifan Lu. (2024). Combined correlation analysis and multilinear regression for strength model of cement-stabilized clayey soils. *International Journal of Geomechanics, 24*(9), Article 04024190 (pp. 01-14).

List of Publications

1. Sonar, N. B., & Pradhan, V. S. (2024, February 7-9). *Determination of Value Indicators for a Real Estate Property.* 7th International Conference on Civil Engineering for Sustainable Development (ICCESD 2024), KUET, Khulna, Bangladesh. (Presented)

- ICCESD 2024 Conference Proceedings (Published)

2. Sonar, N. B., & Pradhan, V. S. (2024, August 23-24). *Ascertaining Valuation Parameters of Real Estate Property.* 8th International Conference on Construction, Real Estate, Infrastructure and Project Management (ICCRIP 2024), NICMAR, Pune, Maharashtra, India. (Presented)

- Taylor and Francis Conference Proceedings (Under review)

3. Sonar, N. B., & Pradhan, V. S. (2024). Effect of Coronavirus Pandemic on Real Estate Property Values. *B. Aadhar International Multidisciplinary Journal, 495* (11), pp. 136-139. (Published – Peer reviewed)

4. Sonar, N. B., & Pradhan, V. S. (2024). Selecting the Crucial Value Indicators for Real Estate Valuation. *NICMAR Journal of Construction Management, 39* (1-4), pp. 05-23. (Published – UGC indexed)

5. Sonar, N. B., & Pradhan, V. S. (2025). Analysing Scenario of Real Estate Properties During the Pandemic 2020. *Journal of Institute of Town Planners, India*. (Accepted – UGC indexed)

6. Sonar, N. B., & Pradhan, V. S. (2025). Influence of Saleable Area on Real Estate Property Values During the 2020 Pandemic. *Journal of Institute of Town Planners, India*. (Accepted – UGC indexed)

7. Sonar, N. B., & Pradhan, V. S. (2025). Predicting Real Estate Property Values Using Machine Learning Techniques. *Journal of Soft Computing in Civil Engineering*. (Under review – SCOPUS indexed)

8. Patent – Republic of South Africa – Patents Act, 1978
   Patentees – Sonar Nandkishor Balaji, Pradhan Vijaya Sitaram
   Title of invention – A Total Weightage Score System for Real Estate Property Valuation
   Granted on 18/12/2024

PATENT REGISTER SHEET

REPUBLIC OF SOUTH AFRICA	REGISTER OF PATENTS	PATENTS ACT, 1978

Official application No.	Lodging date: Provisional	Acceptance date
21 01 2024/03680	22	47 2024/11/22
International classification	**Lodging date: Complete**	**Granted date**
51 G06Q	23 2024/05/13	2024/12/18

71 Full name(s) of applicant(s)/Patentee(s):

SONAR, Nandkishor Balaji
ASSISTANT PROFESSOR, CIVIL ENGINEERING DEPARTMENT OF JAWAHARLAL NEHRU ENGINEERING COLLEGE, MGM (MAHATMA GANDHI MISSION) UNIVERSITY, CHHATRAPATI, SAMBHAJINAGAR, MAHARASHTRA, 431003, India
PRADHAN, Vijaya Sitaram
PROFESSOR, CIVIL ENGINEERING DEPARTMENT OF JAWAHARLAL NEHRU ENGINEERING COLLEGE, MGM, CIVIL ENGINEERING DEPARTMENT OF JAWAHARLAL NEHRU ENGINEERING COLLEGE, MGM (MAHATMA GANDHI MISSION) UNIVERSITY, CHHATRAPATI, SAMBHAJINAGAR, MAHARASHTRA, 431003, India

71 Applicant substitued:	Date registrered

71 Assignee(s):	Date registrered

72 Full name(s) of inventor(s):

SONAR, Nandkishor Balaji
PRADHAN, Vijaya Sitaram

Priority claimed:	Country	Number	Date

54 Title of invention

A TOTAL WEIGHTAGE SCORE SYSTEM FOR REAL ESTATE PROPERTY VALUATION

Address of applicant(s)/patentee(s):

ASSISTANT PROFESSOR, CIVIL ENGINEERING DEPARTMENT OF JAWAHARLAL NEHRU ENGINEERING COLLEGE, MGM (MAHATMA GANDHI MISSION) UNIVERSITY, CHHATRAPATI, SAMBHAJINAGAR, MAHARASHTRA, 431003
INDIA
PROFESSOR, CIVIL ENGINEERING DEPARTMENT OF JAWAHARLAL NEHRU ENGINEERING COLLEGE, MGM, CIVIL ENGINEERING DEPARTMENT OF JAWAHARLAL NEHRU ENGINEERING COLLEGE, MGM (MAHATMA GANDHI MISSION) UNIVERSITY, CHHATRAPATI, SAMBHAJINAGAR, MAHARASHTRA, 431003
INDIA

HISTORY SHEET

Date entry made	Description
2024-05-14	Request for the acceptance of a Patent electronically filed on 13/5/2024, numbered 2024/03680
2024-05-14	Proof reading performed automatically
2024-11-22	Application accepted on 22/11/2024.
2024-12-19	Patent advertised on 18-12-2024.
2024-12-19	Patent granted on 18-12-2024.

Research scholar **Research guide**

N. B. Sonar **Dr. V. S. Pradhan**